

FIGURE 1

GGCATCTGCCCCGAGGAGACCACGCTCCTGGAGCTCTGCTGTCTTCTCAGGGAGACTCTGAGG
CTCTGTTGAGAATCATGCTTTGGAGGCAGCTCATCTATTGGCAACTGCTGGCTTTGTTTTTC
CTCCCTTTTTGCCTGTGTCAAGATGAATACATGGAGTCTCCACAAACCGGAGGACTACCCCC
AGACTGCAGTAAGTGTGTGTCATGGAGACTACAGCTTTCGAGGCTACCAAGGCCCCCCTGGGC
CACCGGGCCCTCCTGGCATTCCAGGAAACCATGGAAACAATGGCAACAATGGAGCCACTGGT
CATGAAGGAGCCAAAGGTGAGAAGGGCGACAAAGGTGACCTGGGGCCTCGAGGGGAGCGGGG
GCAGCATGGCCCCAAAGGAGAGAAGGGCTACCCGGGGATTCCACCAGAACTTCAGATTGCAT
TCATGGCTTCTCTGGCAACCCACTTCAGCAATCAGAACAGTGGGATTATCTTCAGCAGTGTT
GAGACCAACATTGGAAACTTCTTTGATGTCATGACTGGTAGATTTGGGGCCCCAGTATCAGG
TGTGTATTTCTTCACCTTCAGCATGATGAAGCATGAGGATGTTGAGGAAGTGTATGTGTACC
TTATGCACAATGGCAACACAGTCTTCAGCATGTACAGCTATGAAATGAAGGGCAAATCAGAT
ACATCCAGCAATCATGCTGTGCTGAAGCTAGCCAAAGGGGATGAGGTTTGGCTGCGAATGGG
CAATGGCGCTCTCCATGGGGACCACCAACGCTTCTCCACCTTTGCAGGATTCCTGCTCTTTG
AAACTAAGTAAATATATGACTAGAATAGCTCCACTTTGGGGAAGACTTGTAGCTGAGCTGAT
TTGTTACGATCTGAGGAACATTAAAGTTGAGGGTTTTACATTGCTGTATTCAAAAAATTATT
GGTTGCAATGTTGTTACGCTACAGGTACACCAATAATGTTGGACAATTCAGGGGCTCAGAA
GAATCAACCACAAAATAGTCTTCTCAGATGACCTTGACTAATACTCAGCATCTTTATCAC
TCTTTCCTTGGCACCTAAAAGATAATTCTCCTCTGACGCAGGTTGGAAATATTTTTTTCTAT
CACAGAAGTCATTTGCAAAGAATTTTGACTACTCTGCTTTTAATTTAATACCAGTTTTTCAGG
AACCCTGAAGTTTTAAGTTCATTATTCTTTATAACATTTGAGAGAATCGGATGTAGTGATA
TGACAGGGCTGGGGCAAGAACAGGGGCACTAGCTGCCTTATTAGCTAATTTAGTGCCCTCCG
TGTTTCAGCTTAGCCTTTGACCCTTTCCTTTTGATCCACAAAATACATTAAAACTCTGAATTC
ACATACAATGCTATTTTAAAGTCAATAGATTTTAGCTATAAAGTGCTTGACCAGTAATGTGG
TTGTAATTTTGTGTATGTTCCCCCACATCGCCCCAACTTCGGATGTGGGGTCAGGAGGTTG
AGGTTCACTATTAACAAATGTCATAAATATCTCATAGAGGTACAGTGCCAATAGATATTCAA
ATGTTGCATGTTGACCAGAGGGATTTTATATCTGAAGAACATACACTATTAATAAATACCTT
AGAGAAAGATTTTGACCTGGCTTTAGATAAACTGTGGCAAGAAAAATGTAATGAGCAATAT
ATGGAAATAAACACACCTTTGTTAAAGATAAAAAAAA

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44686

><subunit 1 of 1, 246 aa, 1 stop

><MW: 26994, pI: 6.43, NX(S/T): 0

MLWRQLIYWQLLALFFLPFCLCQDEYMESPQTGGLPPDCSKCCHGDYSFRGYQGPPGPPGPP
GIPGNHGNNGNNGATGHEGAKGEKGDGDLGPRGERGQHGPKEGKGYPGIPPELQIAFMASL
ATHFSNQNSGIIFFSSVETNIGNFFDVMGTGRFGAPVSGVYFFTFSMMKHEDVEEVYVYLMHNG
NTVFSMYSYEMKGKSDTSSNHA VLK LAKGDEVWLRMGNGALHGDHQRSTFAGFLLFETK

Important features of the protein:

Signal peptide:

amino acids 1-22

Motif name: Clq domain signature.

amino acids 137-167

Clq domain proteins.

amino acids 135-169, 202-221, 235-244, 57-91, 60-94, 54-88, 81-114, 78-111, 63-96, 51-84, 45-78, 48-81, 33-66, 66-99 and 42-75

FIGURE 3

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGG
CTCTGGCAGGCTCCTGGCAGCATGGCAGTGAAGCTTGGGACCTCCTGCTGGCCCTTGCCCT
GGGCCTGGCCCAGCCAGCCTCTGCCCCGCCGGAAGCTGCTGGTGTCTTCTGCTGGATGGTTTTT
GCTCAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTTTTCAAAGAGATTGTGAGC
AGGGGAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTATTA
TACCCTAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACACATGTGGGACC
CCACCACCAACAAGTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTGG
AATGGATCAGAACCTCTGTGGGTCACTCTGACCAAGGCCAAAAGGAAGGTCTACATGTACTA
CTGGCCAGGCTGTGAGGTTGAGATTCTGGGTGTCAGACCCACCTACTGCCTAGAATATAAAA
ATGTCCCAACGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAAGAGT
GGCCGGGCCGACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGG
GCCTGCATCTCCGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGA
CCAAGTGGATCCAGGAGCGGGGCCTGCAGGACCGCCTGAACGTCATTATTTTCTCGGATCAC
GGAATGACCGACATTTTCTGGATGGACAAAGTGATTGAGCTGAATAAGTACATCAGCCTGAA
TGACCTGCAGCAAGTGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCCTGGGAAAC
ACTCTGAGATATATAACAACTGAGCACAGTGAACACATGACTGTCTACGAGAAAGAAGCC
ATCCCAAGCAGGTTCTATTACAAGAAAGGAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGA
TGAAGGCTGGTTCATAACTGAGAATCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCA
GGCGGGAAGGTTGGCAGCGTGGATGGCACGGCTACGACAACGAGCTCATGGACATGCGGGGC
ATCTTCTGGCCTTCGGACCTGATTTCAAATCCAACCTCAGAGCTGCTCCTATCAGGTCGGT
GGACGTCTACAATGTCATGTGCAATGTGGTGGGCATCACCCCGCTGCCCAACAACGGATCCT
GGTCCAGGGTGATGTGCATGCTGAAGGGCCGCGCCGGCACTGCCCCGCCTGTCTGGCCCAGC
CACTGTGCCCTGGCACTGATTCTTCTCTTCTGCTTGCATAACTGATCATATTGCTTGTCTC
AGAAAAAACACCATCAGCAAAGTGGGCCTCCAAAGCCAGATGATTTTCATTTTATGTGTGA
ATAATAGCTTCATTAACACAATCAAGACCATGCACATTGTAAATACATTATTCTTGGATAAT
TCTATACATAAAAGTTCCTACTTGTTAA

FIGURE 4

MAVKLGTLALLALGLAQPASARRKLLVFLLDGFRSDYISDEALES LPGFKEIVSRGVKVDY
LTPDFPSLSYPNYYTLMTGRHCEVHQMIGNYMWDPTTNKSFDIGVNKDSL MPLWWNGSEPLW
VTLTAKARKVYMYYPGCEVEILGVRPTYCLEYKNVPTDINFANAVSDALDSFKSGRADLAA
IYHERIDVEGHHYGPASPQRKDALKAVDTVLKYMTKWIQERGLQDRNLNVIIFSDHGMTDIFW
MDKVIELNKYISLNDLQQVKDRGPVVS LWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRFYY
KKGK FVSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGIFLAFGP
DFKSNFRAAPIRSVDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALI
LLFLLA

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 100-104, 118-122, 341-345, 404-408

N-myristoylation sites.

amino acids 148-154, 365-371

Amidation site.

amino acids 343-347

FIGURE 5

GCCAGGTGTGCAGGCCGCTCCAAGCCCAGCCTGCCCCGCTGCCGCCACC**ATG**ACGCTCCTCC
CCGGCCTCCTGTTTCTGACCTGGCTGCACACATGCCTGGCCCACCATGACCCCTCCCTCAGG
GGGCACCCCCACAGTCACGGTACCCACACTGCTACTCGGCTGAGGAACTGCCCTCGGCCA
GGCCCCCCCCACACCTGCTGGCTCGAGGTGCCAAGTGGGGGCAGGCTTTGCCTGTAGCCCTGG
TGTCCAGCCTGGAGGCAGCAAGCCACAGGGGGAGGCACGAGAGGCCCTCAGCTACGACCCAG
TGCCCGGTGCTGCGGCCGGAGGAGGTGTTGGAGGCAGACACCCACCAGCGCTCCATCTCACC
CTGGAGATACCGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGT
GCCTGTGCAGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTG
CGGCTGCTCCAGAGCCTGCTGGTGCTGCGCCGCCGGCCCTGCTCCCGCGACGGCTCGGGGCT
CCCCACACCTGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTGGGCTGCACCT
GCGTGCTGCCCCGTTCAGTG**TG**ACCGCCGAGGCCGTGGGGCCCCCTAGACTGGACACGTGTGC
TCCCCAGAGGGCACCCCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCCAACACT
ACCCTTGGGGTCTGGGCATTCCCCGTGTCTGGAGGACAGCCCCCACTGTTCTCCTCATCTC
CAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTGCAGAA
AAGGTGTCACACGGCTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCCTTACCCTA
TCACTGGCCTCAGGCCCCGAGGCTGCCTCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCT
TAAACAATTATTTAAGTGTACGTGTATTATTAACTGATGAACACATCCCCAAA

FIGURE 6

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHELLARGAKWGQAL
PVALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTDEDRYPQKL
AFAECLCRGCIDARTGRETAALNSVRLQLSLLVLRRRRPCSRDGSGGLPTPGAFAFHTEFIHVP
VGCTCVLPRSV

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 112-121

N-myristoylation sites.

amino acids 32-38, 55-61, 133-139

Leucine zipper pattern.

amino acids 3-25

Homologous region to IL-17.

amino acids 99-195

FIGURE 7

CGGCCAGGGCGCCGACAGCCCGACCTCACCAGGAGAAC**ATG**CAGCTCGGCACTGGGCTCCTG
CTGGCCGCCGTCCTGAGCCTGCAGCTGGCTGCAGCCGAAGCCATATGGTGTCAACAGTGCAC
GGGCTTCGGAGGGTGCTCCCATGGATCCAGATGCCTGAGGGACTCCACCCACTGTGTACCA
CTGCCACCCGGGTCTCAGCAACACCGAGGATTTGCCTCTGGTCACCAAGATGTGCCACATA
GGCTGCCCCGATATCCCCAGCCTGGGCCTGGGCCCCTACGTATCCATCGCTTGCTGCCAGAC
CAGCCTCTGCAACCATGAC**TGA**CGGCTGCCCTCCTCCAGGCCCCCGGACGCTCAGCCCCAC
AGCCCCACAGCCTGGCGCCAGGGCTCACGGCCGCCCTCCCTCGAGACTGGCCAGCCCACC
TCTCCCGGCCTCTGCAGCCACCGTCCAGCACCGCTTGTCTTAGGGAAGTCCTGCGTGGAGTC
TTGCCTCAATCTGCTGCCGTCCAAGCCTGGGGCCCATCGTGCCTGCCGCCCCCTTCAGGTCCC
GACCTCCCCACAATAAAATGTGATTGGATCGTGTGGTACAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77623

><subunit 1 of 1, 97 aa, 1 stop

><MW: 10160, pI: 6.56, NX(S/T): 0

MQLGTGLLLLAAVLSLQLAAAEAIWCHQCTGFGGCSHGSRCLRDSTHCVTTATRVLSNTEDLP
LVTKMCHIGCPDIPSLGLGPYVSIACCQTSLCNHD

Important features of the protein:

Signal peptide:

amino acids 1-20

N-myristoylation sites.

amino acids 6-11 and 33-38

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 24-34 and 78-88

FIGURE 9

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAG
CGCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCT
AACTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCG
CGGCACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAG
CTGGGCTCGGGCGGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTC
GCGGGCTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCG
CCGCGATGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCC
TTCTGCCCGCGCTGGTTCAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTA
CAAAATGGCCTACTTCCATGAACTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTT
GTGAGAGTGAGGGAGGAGTCTCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAG
AGCATGTTGCAAAACCTGACAAAACCCGGGACAGGGATTCTGATGGTGATTTCTGGATAGG
GCTTTGGAGGAATGGAGATGGGCAACATCTGGTGCCTGCCCAGATCTCTACCAGTGGTCTG
ATGGAAGCAATTCACAGTACCGAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAAG
TGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCA
GTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTA
ATCCAACAGCCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAAT
GTGGTTGTTACTGAAGCAGGTATAATTCCTAATCTAATTTATGTTGTTATACCAACAATACC
CCTGCTCTTACTGATACTGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAGTA
AAGGAAGAACAACAACTAGTCCAACACAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAA
GAAAGTGGCATGGAAGTATAAATACTCATTGACTTGGTTCCAGAATTTTGTAATTTCTGGATC
TGTATAAGGAATGGCATCAGAACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGCAA
GATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTTTATATGTCTATTATTTT
ATTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCAC
CCAAACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTC
GGGAGTATGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTCATAAGTTGTTATCTA
GTCAATGTAATGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAA
GTGTTTGATAAAAAATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTCAATACAT
GCTCTTTTGATTAAAGAACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAG
TACCATAGAAAAAGTTTGTTTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCA
ATGTCTAGGAAATCTCTCAGAAATAAGAAGCTATTTCAATTAAGTGTGATATAAACCTCCTC
AAACATTTTACTTAGAGGCAAGGATTGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAA
TTATTTTGTAGCTTAAAATTAAACAGATTTTGTAAATAATGTAACCTTTGTTAATAGGTGCATAA
ACACTAATGCAGTCAATTTGAACAAAAGAAGTGACATACACAATATAAATCATATGTCTTCA
CACGTTGCCTATATAATGAGAAGCAGCTCTCTGAGGGTTCTGAAATCAATGTGGTCCCTCTC
TTGCCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGGATTGACACTGGAGGCAGATAGTTGC
AAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTGACTATATTAGTATACAAAGAGG
TCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAGACAAGCACAGCACACAGAC
ATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAAATGGGTTGGAACCCATCAGTGATCG
CATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATGGTGGCTCCTTTCTGTCTTATCTCC
TAGTTTCTTCAATGCTTACGCCTTGTTCTTCAAGAGAAAGTTGTAAGTCTCTGGTCTTCA
TATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTTTCTGGGGGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79230

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30431, pI: 6.79, NX(S/T): 3

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFAADFKHPCYKMAYFHELSSRVSFQEARLACE
SEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDG
SNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINP
TAPVEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKG
RTKTSPNQSTLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89 and 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145
and 212-217

FIGURE 11

GGAGAATGAGAGAGCAGTGAGAGTGGAGTCCGGGGTCTGGTCCGGGGTGGTCTGTCTGCTCCTGGCATGCCCTG
CCACAGCCACTGGGCCCCAAGTTGCTCAGCCTGAAGTAGACACCACCCTGGGTGCTGTGCGAGGCCGGCAGGTGG
GCGTGAAGGGGCACAGACCGCCTTGTGAATGTCTTTCTGGGCATTCCATTTGCCAGCCGCCACTGGGCCCTGACC
GGTCTCAGCCCCACACCCAGCACAGCCCTGGGAGGGTGTGCGGGATGCCAGCACTGCGCCCCAATGTGCCTAC
AAGACGTGGAGAGCATGAACAGCAGCAGATTTGTCTCAACGGAAAAACAGCAGATCTTCTCCGTTTTCAGAGGACT
GCCTGGTCTCAACGTCTATAGCCCAGCTGAGGTCCCCGAGGGTCCGGTAGGCCGGTTCATGGTATGGGTCCATG
GAGGCGCTCTGATAACTGGCGCTGCCACCTCCTACGATGGATCAGCTCTGGCTGCCTATGGGGATGTGGTCTGG
TTACAGTCCAGTACCGCCTTGGGGTCTTGGCTTCTTTCAGCACTGGAGATGAGCATGCACCTGGCAACCAGGGCT
TCCTAGATGTGGTAGCTGCTTTGCGCTGGGTGCAAGAAAACATCGCCCCCTTCGGGGGTGACCTCAACTGTGTCA
CTGTCTTTGGTGGATCTGCCGGTGGGAGCATCATCTCTGGCCTGGTCTGTCCCCAGTGGCTGCAGGGCTGTTCC
ACAGAGCCATCACACAGAGTGGGGTTCATCACCACCCAGGGATCATCGACTCTCACCTTGGCCCCCTAGCTCAGA
AAATCGCAAAACCTTGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAGAAGGAG
AAGAGCTGGTCTTAGCAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCACTGTCTTCCCCAAAA
GCCCCAAGGAACTCCTGAAGGAGAAGCCCTTCCACTCTGTGCCCTTCTCATGGGTGTCAACAACCATGAGTTCA
GCTGGCTCATCCCCAGGGGCTGGGGTCTCCTGGATACAATGGAGCAGATGAGCCGGGAGGACATGCTGGCCATCT
CAACACCCGTCTTGACCACTGTGGATGTGCCCCCTGAGATGATGCCACCGTCATAGATGAATACCTAGGAAGCA
ACTCGGACGCACAAGCCAAATGCCAGGCGTTCCAGGAATTCATGGGTGACGTATTCATCAATGTTCCACCGTCA
GTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTTCTATGAGTTCAGCATCGACCCAGTTCTTTTG
CGAAGATCAAACCTGCCTGGGTGAAGGCTGATCATGGGGCCGAGGGTGCTTTTGTGTTTCGGAGGTCCCTTCCCTCA
TGGACGAGAGCTCCCGCCTGGCCTTTCCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC
AGTGGACCCACTTTGCCCCGACAGGGGACCCCAATAGCAAGGCTCTGCCTCCTTGGCCCCAATTCAACCAGGCGG
AACAATATCTGGAGATCAACCCAGTGCCACGGGCGGACAGAAGTTCAGGGAGGCCTGGATGCAGTTCTGGTCAG
AGACGCTCCCCAGCAAGATAACAAGTGGCACCAGAAGCAGAAGAAGGAGGAGGACCTCTGAGGCGC
AGCCTGAACCTTCTTGGCTGGGGCAAACCACTCTTCAAGTGGTGGCAGAGTCCCAGCACGGCAGCCCGCCTCTC
CCCCTGCTAGGGGGTGTGCTGAGTGAGTGAGTGAAACACAGAATATGGGAATGGCAGCTGCTGAACCTTGAAC
ATGTACAAAGGCCGCTCCACCTCTGGGGCATTGTACAAGTTCTTCCCTCTCCCTGAAGTGCCTTCTGCTTT
CTTCGTGGTAGGTTCTAGCACATTCTCTAGCTTCTTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTC
TGGGCTGTGCGGCCCCGAGTCTGCGTCCATTAGAGCACAGTCCACCCGAGGCTAGCACCGTGTCTGTGTCTGTCT
CCCCCTCAGAGGAGCTCTCTCAAAATGGGGATTAGCCTAACCCCACTCTGTACCCACACCAGGATCGGGTGGGA
CCTGGAGCTAGGGGGTGTGCTGAGTGAGTGAGTGAAACACAGAATATGGGAATGGCAGCTGCTGAACCTTGAAC
CCAGAGCCTTCAGGTGCCAAAGCCATACTCAGGCCCCACCGACATTGTCCACCCTGGCCAGAAGGGTGCATGCC
AATGGCAGAGACCTGGGATGGGAGAAGTCTTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCTGAC
TAAGGCCTCAGACTAGGGCGGGAGGGGTCTCCTCCTCTCTGCTGCCAGTCCCTGGCCCCCTGCACAAGACAACAGA
ATCCATCAGGGCCATGAGTGTACCCAGACCTGACCCTCACCATTCCAGCCCCCTGACCCTCAGGACGCTGGATG
CCAGCTCCCAGCCCCAGTGCCGGGTCTCCCTCCCTTCTTGGCTTGGGGAGACCAGTTTCTGGGGAGCTTCCAAG
AGCACCCACCAAGACACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAGGGCATCCGTCGGGCTATTGTACAC
GAGAAAAGAAGAGACCCACCCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTTCAGATGGAAGTGAGAG
GTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGTCTCTCCCTGCCGCTCTGCCTGGGCTCCCACTTTGGCA
GCACTTGAGGAGCCCTTCAACCCGCGCTGCACTGTAGGAGCCCTTTCTGGGCTGGCCAAGGCCGGAGCCAGCT
CCCTCAGCTTGGCGGGAGGTGCGGAGGGAGAGGGGCGGGCAGGAACCGGGGCTGCGCGCAGCGCTTGGGGCCAG
AGTGAGTTCCGGGTGGGCGTGGGCTCGGCGGGGCCCCACTCAGAGCAGCTGGCCGGCCCCAGGCAGTGAGGGCCT
TAGCACCTGGGCCAGCAGCTGCTGTGCTCGATTCTCGCTGGGCTTAGCTGCCTCCCCGCGGGCAGGGCTCGG
GACCTGCAGCCCTCCATGCCTGACCCTCCCCCACCCTCGTGGGCTCCTGTGCGGCCGAGCCTCCCCAAGGAG
CGCCGCCCCCTGCTCCACAGCGCCAGTCCCATCGACCACCAAGGGCTGAGGAGTGCGGGTGCACAGCGCGGGA
CTGGCAGGCAGCTCCACCTGCTGCCCCAGTGTGGATCCACTGGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGG
GGACTTGAGAACTTTATGTCTAGCTAAGGGATTGTAAATACACCGATGGGCACTCTGTATCTAGCTCAAGGTT
TGTAACACACCAATCAGCACCTGTGTCTAGCTCAGTGTGTGTGAATGCACCAATCCACACTCTGTATCTGGCT
ACTCTGGTGGGGACTTGAGAACTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGAGAACCT
TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGCACCTGTCAAACAGACCACTTGACTCTCTGTAAAT
GGACCAATCAGCAGGATGTGGGTGGGGCGAGACAAGAGAATAAAGCAGGCTGCCTGAGCCAGCAGTGACAACCC
CCCTCGGGTCCCCTCCCACGCCGTGGAAGCTTTGTTCTTTTCGCTCTTTGCAATAAATCTTGCTACTGCCAAAA

FIGURE 12

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79862

><subunit 1 of 1, 571 aa, 1 stop

><MW: 62282, pI: 5.56, NX(S/T): 1

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVFLG
IPFAQPPLGPDRFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDC
LVLNVYSPAIEVPAGSGRPVMVWVHGGALITGAATSYDGSALAAYGDVVVTVQYRLGVLGFF
STGDEHAPGNQGFLDVVAALRWVQENIAPFGGDLNCVTVFGGSAGGSIISGLVLSPVAAGLF
HRAITQSGVITTPGIIDSHPWPLAQKIAN TLACSSSSPAEMVQCLOQKEGEELVLSKKLKNT
IYPLTVDGTVPKSPKELLKEKPFHSPFLMGVNNHEFSWLI PRGWGLLD TMEQMSREDMLA
ISTPVLTS LDVPPMMPTVIDEYLG SNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDGS
PVFFYEFQHRPSSFAKIKPAWVKADHGAEGAFVFGGPF LMDESSRLAFPEATEEEKQLSLTM
MAQWTHFARTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLP SKIQW
HQKQKNRKAQEDL

Important features of the protein:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 226-245

N-glycosylation site.

amino acids 105-109

N-myristoylation sites.

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,
162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,
461-467

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

Carboxylesterases type-B serine active site.

amino acids 216-232

FIGURE 13

CATGGAGCCTCTTGCAGCTTACCCGCTAAATGTTCCGGGCCCAGAGCAAAGGTATTTGCAG
TTTTGCTGTCTATAGTTCTATGCACAGTAACGCTATTTCTTCTACAACATAAAATTCCTCAA
CCTAAAATCAACAGCTTTTATGCCTTTGAAGTGAAGGATGCAAAAGGAAGAACTGTTTCTCT
GGAAAAGTATAAAGGCAAAGTTTCACTAGTTGTAAACGTGGCCAGTGACTGCCAACTCACAG
ACAGAAATTACTTAGGGCTGAAGGAACTGCACAAAGAGTTTGGACCATCCCACCTCAGCGTG
TTGGCTTTTCCCTGCAATCAGTTTGGAGAATCGGAGCCCCGCCAAGCAAGGAAGTAGAATC
TTTTGCAAGAAAAAACTACGGAGTAACTTTCCCCATCTTCCACAAGATTAAGATTCTAGGAT
CTGAAGGAGAACCCTGCATTTAGATTTCTTGTTGATTCTTCAAAGAAGGAACCAAGGTGGAAT
TTTTGGAAGTATCTTGTCAACCCTGAGGGTCAAGTTGTGAAGTTCTGGAGGCCAGAGGAGCC
CATTGAAGTCATCAGGCCTGACATAGCAGCTCTGGTTAGACAAGTGATCATAAAAAAGAAAG
AGGATCTA**TGA**GAAATGCCATTGCGTTTCTAATAGAACAGAGAAATGTCTCCATGAGGGTTTG
GTCTCATTTTAAACATTTTTTTTTTTGGAGACAGTGTCTCACTCTGTCACCCAGGCTGGAGTG
CAGTAGTGC GTTCTCAGCTCATTGCAACCTCTGCCTTTTTTAAACATGCTATTAAATGTGGCA
ATGAAGGATTTTTTTTTTAATGTTATCTTGCTATTAAGTGGTAATGAATGTTCCCAGGATGAG
GATGTTACCCAAAGCAAAAATCAAGAGTAGCCAAAGAATCAACATGAAATATATTAACACT
TCCTCTGACCATACTAAAGAATTCAGAATACACAGTGACCAATGTGCCTCAATATCTTATTG
TTCAACTTGACATTTTCTAGGACTGTACTTGATGAAAATGCCAACACACTAGACCACTCTTT
GGATTCAAGAGCACTGTGTATGACTGAAATTTCTGGAATAACTGTAAATGGTTATGTTAATG
GAATAAAACACAAATGTTGAAAAATGTAAATATATATACATAGATTCAAATCCTTATATAT
GTATGCTTGTTTTGTGTACAGGATTTTGTCTTTTCTTTTAAAGTACAGGTTCTAGTGTTTT
ACTATAACTGTCACATATGTATGTAAGTACATATATAAATAGTCATTTATAAATGACCGTAT
TATAACATTTGAAAAAGTCTTCATCAAAAAAAAAAAAAA

FIGURE 14

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80136

><subunit 1 of 1, 209 aa, 1 stop

><MW: 23909, pI: 9.68, NX(S/T): 0

MEPLAAYPLKCSGPRAKVFAVLLSIVLCTVTLFLLQLKFLKPKINSFYAFEVKDAKGRTVSL
EKYKGKVS�VVNVASDCQLTDRNYLGLKELHKEFGPSHFVLAFFPCNQFGESEPRPSKEVES
FARKNYGVTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNFWKYLVNPEGQVVKFWRPEEP
IEVIRPDIAALVRQVVIKKKEDL

Important features of the protein:

Signal peptide:

amino acids 1-31

Glutathione peroxidases signature 2.

amino acids 104-112

Glutathione peroxidases.

amino acids 57-82

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 15

TGTCGCCTGGCCCTCGCCATGCAGACCCCGCGAGCGTCCCCTCCCCGCCCGGCCCTCCTGCTTCTGCTGCTGCTA
CTGGGGGGGCGCCACGGCCTCTTTCTGAGGAGCGCGCGCGCTTAGCGTGGCCCCCAGGGACTACCTGAACAC
TATCCCGTGTGTTGTGGGCAGCGGGCCCGGACGCCTGACCCCGCAGAAGGTGCTGACGACCTCAACATCCAGCGA
GTCTTGCGGGTCAACAGGACGCTGTTTCATTGGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCAGC
TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAACGTGTGTGCGGATGAAG
GGCAAACAGGAGGGCGAGTGTCTGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC
GGTTCCAACGCCTTCAACCCGGTGTGCGCCAACCTACAGCATAGACACCCCTGCAGCCCGTGGAGACAACATCAGC
GGTATGGCCCGCTGCCCCGTACGACCCCAAGCACGCCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTCACAGCT
ACTGTTACCGACTTCTTAGCCATTGATGCTGTCTATCTACCGCAGCCTCGGGGACAGGCCCCACCTGCGCACCGTG
AAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC
TTCCGGGAGATTGCGATGGAGTTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCCGAGTGTGCAAGAAC
GACGTGGGAGGCTCCCCCGCGTGTGGAGAAGCAGTGGACGTCCTTCTGAAGGCGCGGCTCAACTGCTCTGTA
CCCGGAGACTCCCATTCTTCAACGTGCTGCAGGCTGTACGGGCGTGGTCAGCCTCGGGGGCGGCCCGCTG
GTCTGGCCGTTTTTTTCCACGCCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG
GCAGCTGTGTTTGAAGGCCGCTTCCGAGAGCAGAAGTCCCCGAGTCCATCTGGACGCCGGTGGCGGAGGATCAG
GTGCCTCGACCCCGGCCCGGGTGTGCGCAGCCCCCGGGATGCAGTACAAATGCTCCAGCGCCTTGCCGGATGAC
ATCCTCAACTTTGTCAAGACCCACCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCCTG
CGGACCTGATGAGGACAGCTGACTCGAGTGGCTGTGGACGTGGGAGCCGGCCCCCTGGGGCAACCAGACCGTT
GTCTTCTCGGGTCTGAGGCGGGGACGTCCTCAAGTTCCCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGCTC
GGGCTCAGTGTCTTCTGGAGGAGTTTGAACCTACCGGCCGGACAGGTGTGGACGGCCCCGGGGGCGGAGACA
GGGCAGCGGCTGCTGAGCTTGGAGCTGGACGAGCTTCGGGGGGCCTGCTGGCTGCCTTCCCCCGCTGCGTGGT
CGAGTGCCTGTGGCTCGCTGCCAGCACTCTCGGGGTGTATGAAGAACTGTATCGGCAGTCAAGACCCCTACTGC
GGGTGGGCCCCCGACGGCTCCTGCATCTTCTCAGCCCGGCCACCAAGAGCCGCTTTGAGCAGGACGTGTCCGGG
GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCGAGGACCGCGCGGGGCTGGT
TCGGTGAACCTGCTGTTAACCTCGTCCGTGGCGGCCCTTCGTGGTGGGAGCCGTGGTGTCCGGCTTCAAGCTGGG
TGGTTCTGTTGGGCTCCGTGAGCGGGGAGCTGGCCCGGCGCAAGGACAAGGAGGCCATCCTGGCGCACGGGGCG
GGCGAGGCGGTGCTGAGCGTACGCCCTGGGCGAGCGCAGGGCGCAGGGTCCCGGGGGCGGGGGCGGAGGCGGT
GGCGGTGGCGCCGGGGTTCCCCCGGAGGCCCTGCTGGCGCCCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG
CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGCTGCTGCCACGCCCCGAGCAGACGCCGCTGCCGCAAGCGC
CTGCCCCATCCGCACCCGCACCCCTGGGCCCCCGCGCTGGGACCACGGCCACCCCTGCTCCCGGCC
TCCGCTTCATCCTCCTCCTGCTGTGGCGCCCCCGCGGCCCGGAGCAGCCCCCGCGCCTGGGGAGCCGACC
CCGACGGCCGCTCTATGCTGCCCGGCCCGGCCGCGCCTCCACGGCGACTTCCCGCTCACCCCCACGCCAGC
CCGACCGCCGGCGGGTGGTGTCCGCGCCACGGGCCCTTGGACCCAGCCTCAGCCGCGGATGGCCTCCCGCGG
CCCTGGAGCCCGCCCCGACGGGCAGCTGAGGAGGCCACTGGGCCCCCAGCCCCCTCCGGCCGCCACCTGCGC
CGCACCCACAGTTCAACAGCGGCGAGGCCCGGCCCTGGGGACCGCCACCGCGGCTGCCACGCCCGGCCGGGACA
GACTTGGCCCACTCCTCCCCCTATGGGGGGGCGGACAGGACTGCGCCCCCGTGGGCTAGGCGGGGGCCCCCG
ATGCCCTTGGCAGTGCCAGCCACGGGAACCAGGAGCGAGAGACGTTGCCAGAACGCCGGGGCCCCGGGGCAACTCCG
AGTGGGTGCTCAAGTCCCCCGCGACCCACCCGCGGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACAACAG
CTCGCCCTCCCCCTACCCGGGGCGCAGGACGCTGAGACGTTTGGGGGTGGGTGGGCGGGAGGACTTTGCTATG
GATTTGAGGTTGACCTTATGCGCGTAGGTTTTTGGTTTTTTTTTTCAGTTTGGTTTTCTTTTTCGCGTTTTCTAAC
AATTGCACAACCTCCGTTCTCGGGGTGGCGGCAGGCAGGGGAGGCTTGGACGCCGGTGGGGAATGGGGGGCCACAG
CTGCAGACCTAAGCCCTCCCCCACCCTGGAAAGTCCCTCCCCAACCCAGGCCCTTGGCGTGTGTGGGTGTGCG
TGCGTGTGCGTGCCGTGTTCGTGTGCAAGGGGCCGGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTG
TGGGCGTGTGTGTCAAGTGGGCCACGCGTGCAGGGTGTGTGTCCAGAGCGACGATCGTGGTGGCCCCAGCGGCC
TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGTCTCCGAGGTGCCGGTTAGGAGTTTGAAC
CCCCCCTCTGCAAGGGGAAGCGGGGACAATGCCGGGGTTTCAAGCAGGAGACACGAGGAGGGCCTGCCCGGA
AGTCACATCGGCAGAGTGTCTAAAGGGCTTGGGGGCTGGGGGGCGGCAAGGTGGGTGGGGCCCCCTCTGTAA
ATACGGCCCCAGGGTGGTGAGAGAGTCCCATGCCACCCGTCCCCTTGTGACCTCCCCCTATGACCTCCAGCTGA
CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTTGGAGTTTGCCTCCCCAGCCCCCTCCCCATCAAT
AAAACCTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 16

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80145

><subunit 1 of 1, 888 aa, 1 stop

><MW: 95285, pI: 8.89, NX(S/T): 8

MQTPRASPPRPALLLLLLLLGGAHGLFPEEPPPLSVAPRDYLNHYPVFVVGSGPGRLTPAEGA
DDLNIQRVLRVNRTLFIGDRDNLRYVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEG
ECRNFVKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALF
SDGMLFTATVTDFLAIDAVIYRSLGDRPTLRRTVKHDSKWFKEPYFVHAVEWGS HVYFFFREI
AMEFNYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCVPGDSHFYFNVLQAVTG
VVS LGGRPVLAVFSTPSNSIPGSAVCAFDLTQVA AVFEGRFREQKSPESI WTPVPEDQVPR
PRPGCCAAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHPWILRTL MRHQLTRVAVDV
GAGPWGNQTVVFLGSEAGTVLKF LVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGGETGQ
RLLSLELDAASGGLLA AFPRCVVRVPVARCQQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPG
TRAAFEQDVSGASTSG LGDCTG LLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGW
FVGLRERRELARRKDKEAILAHGAGEAVLSVSR LGERRAQGP GGRGGGGGGGAGVPPEALLA
PLMQNGWAKATLLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHPHALGPRAWDHGHPLLLPA
SASSSLLLLAPARAPEQPPAPGEPTPDGRLYAARPGRASHGDFPLTPHASPDRRRRVVSAPTG
PLDPASAADGLPRPWSPPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPG
TDLAHL LPPYGGADRTAPPVP

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 318-339, 598-617

N-glycosylation sites.

amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445,
462-466

Glycosaminoglycan attachment sites.

amino acids 51-55, 573-577

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation sites.

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,
490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,
574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,
668-674, 669-675, 670-676, 868-874, 879-885

FIGURE 17

AGCAACTCAAGTTCATCATTGTCCTGAGAGAGAGGAGCAGCGCGTTCTCGGCCGGGACAGC
AGAACGCCAGGGGACCCTCACCTGGGCGCGCCGGGGCACGGGCTTTGATTGTCCTGGGGTCG
CGGAGACCCGCGCGCCTGCCCTGCACGCCGGGCGGCAACCTTTGCAGTCGCGTTGGCTGCTG
CGATCGGCCGGCGGGTCCCTGCCGAAGGCTCGGCTGCTTCTGTCCACCTCTTACACTTCTTC
ATTTATCGGTGGATCATTTCGAGAGTCCGTCTTGTAATGTTTGGCACTTTGCTACTTTATT
GCTTCTTTCTGGCGACAGTTCCAGCACTCGCCGAGACCGGCGGAGAAAGGCAGCTGAGCCCCG
GAGAAGAGCGAAATATGGGGACCCGGGCTAAAAGCAGACGTCGTCCTTCCCGCCCGCTATTT
CTATATTCAGGCAGTGGATACATCAGGGAATAAATTACATCTTCTCCAGGCGAAAAGGTCT
TCCAGGTGAAAGTCTCAGCACCAGAGGAGCAATTCAGTAGAGTTGGAGTCCAGGTTTTAGAC
CGAAAAGATGGGTCCTTCATAGTAAGATACAGAATGTATGCAAGCTACAAAAATCTGAAGGT
GGAAATTAAATTCCAAGGGCAACATGTGGCCAAATCCCCATATATTTTAAAAGGGCCGGTTT
ACCATGAGAACTGTGACTGTCCTCTGCAAGATAGTGCAGCCTGGCTACGGGAGATGAACTGC
CCTGAAACCATTGCTCAGATTCAGAGAGATCTGGCACATTTCCCTGCTGTGGATCCAGAAAA
GATTGCAGTAGAAATCCCCAAAAGATTTGGACAGAGGCAGAGCCTATGTCACTACACCTTAA
AGGATAACAAGGTTTATATCAAGACTCATGGTGAACATGTAGGTTTTAGAATTTTCATGGAT
GCCATACTACTTTCTTTGACTAGAAAGGTGAAGATGCCAGATGTGGAGCTCTTTGTTAATTT
GGGAGACTGGCCTTTGGAAAAAAGAAATCCAATTCAAACATCCATCCGATCTTTTCCTGGT
GTGGCTCCACAGATTCCAAGGATATCGTGATGCCTACGTACGATTTGACTGATTCTGTTCTG
GAAACCATGGGCCGGGTAAAGTCTGGATATGATGTCCGTGCAAGCTAACACGGGTCCTCCCTG
GGAAAGCAAAAATTCAGTCCGTCTGGAGAGGGCGAGACAGCCGCAAAGAGAGACTCGAGC
TGGTTAACTCAGTAGAAAACACCCAGAACTCATAGACGCTGCTTTCACCAACTTTTTCTTC
TTTAAACACGATGAAAACCTGTATGGTCCCATTTGTGAAACATATTTTCAATTTTTTGATTTCTT
CAAGCATAAGTATCAAATAAATATCGATGGCACTGTAGCAGCTTATCGCCTGCCATATTTGC
TAGTTGGTGACAGTGTTGTGCTGAAGCAGGATTCCATCTACTATGAACATTTTTACAATGAG
CTGCAGCCCTGGAAACACTACATTCCAGTTAAGAGCAACCTGAGCGATCTGCTAGAAAAACT
TAAATGGGCGAAAGATCACGATGAAGAGGCCAAAAAGATAGCAAAAGCAGGACAAGAATTTG
CAAGAAATAATCTCATGGGCGATGACATATTCTGTTATTATTTCAAACTTTTCCAGGAATAT
GCCAATTTACAAGTGAGTGAGCCCCAAATCCGAGAGGGCATGAAAAGGGTAGAACCACAGAC
TGAGGACGACCTCTTCCCTTGTACTTGCCATAGGAAAAAGACCAGATGAACTCTGATATG
CAAAATAACTTCTATTAGAATAATGGTGCTCTGAAGACTCTTCTTAACTAAAAAGAAGATT
TTTTTAAGTATTAATTCATGGACAATATAAAATCTGTGTGATTGTTTGCAGTATGAAGACA
CATTTCTACTTATGCAGTATTCTCATGACTGTACTTTAAAGTACATTTTLAGAATTTTATAA
TAAACCACCTTTATTTTAAAGGAAAAAA

FIGURE 18

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84917

><subunit 1 of 1, 502 aa, 1 stop

><MW: 58043, pI: 7.94, NX(S/T): 2

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVVLPAFYFYIQAVDTSGNKF
TSSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVEIKFQGQHVAKS
PYILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQORDLAHFPAVDPEKIAVEIPKRFGQR
QSLCHYTLKDNKVYIKTHGEHVGFRIFMDAILLSLTRKVKMPDVELFVNLGDWPLEKKKSNS
NIHPIFSWCGSTDSKDIVMPTYDLTDSVLETMGRVSLDMMSVQANTGPPWESKNSTAVWRGR
DSRKERLELVKLSRKHPELIDAAFTNFFFFKHDENLYGPIVKHISFFDFFKHKYQINIDGTV
AAYRLPYLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKSNLSDLLEKLKWAKDHDEEAKK
IAKAGQEFARNNLMGDDIFCYYFKLFQEYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRK
KTKDEL

Important features of the protein:

Signal peptide:

amino acids 1-17

N-glycosylation sites.

amino acids 302-306, 414-418

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 243-247, 495-499

Tyrosine kinase phosphorylation site.

amino acids 341-348

N-myristoylation sites.

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

Endoplasmic reticulum targeting sequence.

amino acids 499-504

FIGURE 19

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCA
GCCTAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCG
AGGTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCTCTGCC
AGAGCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTG
GACCACGCTGCCACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAG
ATGATTTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAACTTACCTTGTGGAA
AAAGGAGCATCTCACAGTTTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAAT
CAAGGCATTGATGACGGGGAGCCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTC
CTGCACTGCTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTT
TATGGAGATGAAACCTGGGTAAATTAATCCCAAAGCATTTTGTGGAATATGATGGAACAAC
CTCATTTTTCTGTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTTGGATAAAG
TATTAAAAAGAGGAGATTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGC
CACATTTCAGGGCCCCAACAGCCCCCTGATTGGGCAGAAGCTGAGCGAGATGGACAGCGTGCT
GATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGAGACGCCTTTACCCAATTTGCTGG
TTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACGGGGCCTCCTCCACCGAGGAG
GTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAACCCGGTGATATCCGACA
TCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGCACTTGGCTTACCGA
TTCCAAAAGACAGTGTAGGGAGCCTCCTATTCAGTTGTGGAAGGAAGACCAATGAGAGAG
CAGTTGAGATTTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAGAATGTGCC
GTCATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTCAGAAAGATTGCATGGGAACT
GGATCAGACTGTACTTGGAGGAAAAGCATTGAGAGTCTTCAACCTGGGCTCCAAGGTT
CTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCCA
GTTCTCACCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGT
CCCCTGTCTCTCTGGGTTTTCTCTGCTCTTTATTTGGTGATCCTGGTTCTTTCGGCCG
TTCACGTCATTGTGTGCACCTCAGCTGAAAGTTCTGTGCTACTTCTGTGGCCTCTCGTGGCTG
GCGGCAGGCTGCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGG
CAGTGCCCTGGACAGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGG
TGTCCCGACACAGGTGTTACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGGAAGCT
AGGTTCTGCGACTGTTACCAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCC
CCAGCTGAGGGGGTGTGTGAATCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGA
GGGAAGAAGAGACAATCGGCCTGGACACTCAGGAGGGTCAAAGGAGACTTGGTGCACCAC
TCATCCTGCCACCCCCAGAATGCATCCTGCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGA
CGTTTTCTGTTGGAATTCTTAGTCTTGGCCTCGGACACCTTCATTGTTAGCTGGGGAGTG
GTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCACACTCAGATCCACAGAGCCCAGGATCAAG
GGACCACTGCAGTGGCAGCAGGACTGTTGGGCCCCCACCCTGCACAGCCCTCATC
CCCTCTTGGCTTGAGCCGTGAGAGCCCTGTGCTGAGTGTCTGACCGAGACACTCACAGCTT
TGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTGCCACGCTTGCACCTCGGGC
CCATCTGGGCTCATGCTCTCTCTGCTATTGAATTAGTACCTAGCTGCACACAGTATGTA
GTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84920

><subunit 1 of 1, 310 aa, 1 stop

><MW: 33875, pI: 7.08, NX(S/T): 2

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPP
PLFSKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMT
GSLPGFVDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVS
DYTEVDNNVTRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHT
SLQSKERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 21

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATG**GCAAT
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCCTTTGACCTCAGGGCCATG
AAGCTGCCCAGCATCTCCTGTCTCATGAGTGCTTTGAGGCTATCCTGTCTCTGGACACCGG
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGA
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACTGAC
AAATGCAACGCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC
GCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGGCCCTCCTG
CACCACCGAGGGCACCACCAGCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTCACCAGTGCTTCAGCCACCACCCCT
CCCCGAGCACTACAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC
ATAGACCGCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTCTTGCTGCTTCAGC
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576

><subunit 1 of 1, 251 aa, 1 stop

><MW: 26935, pI: 7.42, NX(S/T): 2

MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQSGRMTVGNFSVPVYIRTCHR
PSCCTTEGTTSPWTAIDLQGSCCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVGLSA

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

FIGURE 23

CCCACGCGTCCGGGACAGATGAACTTAAAAGAGAAGCTTTAGCTGCCAAAGATTGGGAAAGG
GAAAGGACAAAAAAGACCCCTGGGCTACACGGCGTAGGTGCAGGGTTTCCTACTGCTGTTCT
TTTATGCTGGGAGCTGTGGCTGTAACCAACTAGGAAATAACGTATGCAGCAGCT**ATG**GCTGT
CAGAGAGTTGTGCTTCCCAAGACAAAGGCAAGTCCTGTTTCTTTTTCTTTTTTGGGGAGTGT
CCTTGGCAGGTTCTGGGTTTGGACGTTATTCGGTGACTGAGGAAACAGAGAAAGGATCCTTT
GTGGTCAATCTGGCAAAGGATCTGGGACTAGCAGAGGGGGAGCTGGCTGCAAGGGGAACCAG
GGTGGTTTCCGATGATAACAAACAATACCTGCTCCTGGATTACATACCGGGAATTTGCTCA
CAAATGAGAACTGGACCGAGAGAAGCTGTGTGGCCCTAAAGAGCCCTGTATGCTGTATTTC
CAAATTTTAATGGATGATCCCTTTCAGATTTACCGGGCTGAGCTGAGAGTCAGGGATATAAA
TGATCACGCGCCAGTATTTTCAAGACAAAGAACAGTCTTAAAAATATCAGAAAATACAGCTG
AAGGGACAGCATTTAGACTAGAAAGAGCACAGGATCCAGATGGAGGACTTAACGGTATCCAA
AACTACACGATCAGCCCCAACTCTTTTTTCCATATTAACATTAGTGGCGGTGATGAAGGCAT
GATATATCCAGAGCTAGTGTGGACAAAGCACTGGATCGGGAGGAGCAGGGAGAGCTCAGCT
TAACCCTCACAGCGCTGGATGGTGGGTCTCCATCCAGGTCTGGGACCTCTACTGTACGCATC
GTTGTCTTGGACGTCAATGACAATGCCCCACAGTTTGGCCAGGCTCTGTATGAGACCCAGGC
TCCAGAAAACAGCCCCATTGGGTTCCCTTATTGTTAAGGTATGGGCAGAAGATGTAGACTCTG
GAGTCAACGCGGAAGTATCCTATTCTTTTTTGATGCCTCAGAAAATATTCGAACGACCTTT
CAAATCAATCCTTTTTCTGGGGAAATCTTTCTCAGAGAATTGCTTGATTATGAGTTAGTAAA
TTCTTACAAAATAAATATACAGGCAATGGACGGTGGAGGCCTTTCTGCAAGATGTAGGGTTT
TAGTGGAAGTATTGGACACCAATGACAATCCCCCTGAACTGATCGTATCATCATTTTCCAAC
TCTGTTGCTGAGAATTCTCCTGAGACGCCGCTGGCTGTTTTTAAGATTAATGACAGAGACTC
TGGAGAAAATGGAAAGATGGTTTGTCTACATTCAAGAGAATCTGCCATTCTACTAAAACCTT
CTGTGGAGAATTTTACATCCTAATTACAGAAGGCGCGCTGGACAGAGAGATCAGAGCCGAG
TACAACATCACTATCACCGTCACTGACTTGGGGACACCCAGGCTGAAAACCGAGCACAAAT
AACGGTCCTGGTCTCCGACGTCAATGACAACGCCCCCGCCTTCACCCAAACCTCTACACCC
TGTTTCGTCCGCGAGAACAACAGCCCCGCCCTGCACATCGGCAGCGTCAGCGCCACAGACAGA
GACTCGGGCACCAACGCCCCAGGTACCTACTCGCTGCTGCCGCCCAAGACCCGCACCTGCC
CCTCGCCTCCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCCGCCCTCAGGTCGCTGG
ACTACGAGGCCCTGCAGGCTTTCGAGTTCGCGCTGGGCGCCACAGACCGCGGCTCCCCCGCG
CTGAGCAGAGAGGCGCTGGTGC GCGTGTGCTGGTGTGCTGGACGCCAACGACAACCTCGCCCTTCGT
GCTGTACCCGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGC
CGGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTG
TCGTACCAGCTGCTCAAGGCCACGGAGCCCCGGGCTGTTGCGTGTGTGGGCGCACAAATGGGGA
GGTGCGCACCGCCAGGCTGCTGAGCGAGCGCGACGACGCCAAGCACAGGCTCGTGGTGTGTTG
TCAAGGACAATGGCGAGCCTCCTCGCTCGGCCACCGCCACGCTGCACTTGCTCCTGGTGGAC
GGCTTCTCCCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGGCCAGGCCAGGCCGAGGC
CGACTTGCTCACCGTCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCCTCT
CGGTGCTCCTGTTTCGTGGCGGTGCGGCTGTGCAGGAGGAGCAGGGCGGCCTCGGTGGGTGCG
TGCTCGGTGCCCCGAGGTCCTTTTTCCAGGGCATCTGGTGGACGTGAGGGGCGCTGAGACCCT
GTCCCAGAGCTACCAGTATGAGGTGTGTCTGACGGGAGGCCCCGGGACCAGTGAGTTCAAGT
TCTTGAAACCAGTTATTTTCGGATATTCAGGCACAGGGCCCTGGGAGGAAGGGTGAAGAAAAT
TCCACCTTCCGAAATAGCTTTGGATTTAATATT**CAGTAA**AGTCTGTTTTTAGTTTCATATAC
TTTTGGTGTGTACATAGCCATGTTTCTATTAGTTTACTTTTAAATCTCAAATTTAAGTTAT
TATGCAACTTCAAGCATTATTTTCAAGTAGTATAACCCCTGTGGTTTTACAATGTTTCATCAT
TTTTTTGCATTAATAACAACCTGGGTTTAAATTTAATGAGTATTTTTTTCTAAATGATAGTGTT
AAGGTTTTAATTCCTTCCAACCTGCCAAGGAATTAATTACTATTATATCTCATTACAGAAAT
CTGAGGTTTTGATTCAATTTCAAGAGCTTGCATCTCATGATTCTAATCACTTCTGTCTATAGTG
TACTTGCTCTATTTAAGAAGGCATATCTACATTTCCAAACTCATTCTAACATTCTATATATT
CGTGTTTGAAAACCATGTCAATTTATTTCTACATCATGTATTTAAAAAGAAATATTTCTCTAC
TACTATGCTCATGACAAAATGAAACAAAGCATATTGTGAGCAATACTGAACATCAATAATAC
CCTTAGTTTTATATACTTATTATTTTATCTTTAAGCATGCTACTTTTACTTGGCCAATATTTT
CTTATGTTAACTTTTGCTGATGTATAAAACAGACTATGCCTTATAATTGAAATAAAATTATA
ATCTGCCTGAAAATGAATAAAAATAAAACATTTTGAAATGTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA87976

><subunit 1 of 1, 800 aa, 1 stop

><MW: 87621, pI: 4.77, NX(S/T): 7

MAVRELCFPRQRQVLFLEFWGVSLAGSGFGGRYSVTEETEEKGSFVVNLAKDLGLAEGELAAR
GTRVVSDDNKQYLLLDSTGNLLTNEKLDREKLCGPKEPCMLYFQILMDDPFQIYRAELRVR
DINDHAPVFQDKETVLKISENTAEGTAFRLERAQDPDGGLNGIQNYTISPNSFFHINISGGD
EGMIYPELVLDKALDREEQGELSITLTALDGGSPSRSGTSTVRIVVLDVNDNAPQFAQALYE
TQAPENSPIGFLIVKVAEDVDSGVNAEVSYSFFDASENIRTTFQINPFSGEIFLRELLDYE
LVNSYKINIQAMDGGGLSARCRVLVEVLDTNDNPPELIVSSFSNSVAENSPETPLAVFKIND
RDSGENGMVCYIQENLPFLKPSVENFYILITEGALDREIRAENITITVTDLGTPRLKTE
HNITVLVSDVNDNAPAFQTQSYTLFVRENNSPALHIGSVSATDRDSGTNAQVTYSLLPPQDP
HLPLASLV SINADNGHLFALRSLDYEALQAFEFVVGATDRGSPALSREALVRVLVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVAVDGDGSGQNAWLSYQLLKATEPGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHLLLVDGFSQPYLPLPEAAPAAQ
AEADLLTVYLVVALASVSSLFLLSVLLFVAVRLCRRSRAASVGRCSVPEGPFPGHLVDVRGA
ETLSQSYQYEVCLTGGPGTSEFKFLKPVISDIQAQGPGRKGEENSTFRNSFGFNIQ

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 687-711

N-glycosylation sites.

amino acids 169-173, 181-185, 418-422, 436-440, 567-571, 788-792

Glycosaminoglycan attachment site.

amino acids 28-32

Tyrosine kinase phosphorylation sites.

amino acids 394-402, 578-585

N-myristoylation sites.

amino acids 22-28, 27-33, 53-59, 82-88, 162-168, 184-190,
217-223, 324-330, 325-331, 471-477, 568-574, 759-765

Amidation site.

amino acids 781-785

Aminoacyl-transfer RNA synthetases class-II signature 1.

amino acids 117-138

Cadherins extracellular repeated domain signature.

amino acids 121-132, 230-241, 335-346, 439-450, 549-560

3.

[illegible]

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92234

><subunit 1 of 1, 507 aa, 1 stop

><MW: 56692, pI: 5.22, NX(S/T): 3

MDPKLGRMAASLLAVLLLLLLERGMFSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIESD
SVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGSDPTK
GTVCFYGHLDVQPADRGDGLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD
LPVNIKFIIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAITYGTRGNSY
FMVEVKCRDQDFHSGTFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEEINTY
KAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSSLIHGIEGAFDEPGTKTVIPGRVIGK
FSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAI
RTVFGTEPDMIRDGSTIPIAKMFQEIVHKSVVLIPLGAVDDGEHSQNEKINRWNYIEGTKLF
AAFFLEMAQLH

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

273-292

N-glycosylation sites.

amino acids 322-326, 382-386, 402-406

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 400-404

N-myristoylation sites.

amino acids 89-95, 119-125, 162-168, 197-203, 242-248, 263-269,
351-357

Cell attachment sequence.

amino acids 140-143

ArgE / dapE / ACY1 / CPG:

amino acids 156-167

FIGURE 27

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCCTGGGAATTTAAGGGACCCACACTACCTTCC
CGAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTC**ATG**GGACCTGTGCGGTT
GGGAATATTGCTTTTCCTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGAAGGAGG
AGGACGATGACACAGAACGCTTGCCCAGCAAATGCGAAGTGTGTAAGCTGCTGAGCACAGAG
CTACAGGCGGAACTGAGTCGCACCGGTGATCTCGAGAGGTGCTGGAGCTGGGGCAGGTGCT
GGATACAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAAGGCTGGAAGAGG
CCTTAGAGAATTTATGTGAGCGGATCCTGGACTATAGTGTTACGCTGAGCGCAAGGGCTCA
CTGAGATATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAGGCCTAGTGCGAAGGG
GGTGAAGGTGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGCCCAGCGTGGAGGTCACAT
ACCTCAAGAAGCAGTGTGAGACCATGTTGGAGGAGTTTGAAGACATTGTGGGAGACTGGTAC
TTCCACCATCAGGAGCAGCCCCTACAAAATTTTCTCTGTGAAGGTCATGTGCTCCCAGCTGC
TGAAACTGCATGTCTACAGGAACTTGGACTGGAAAGGAGATCACAGATGGGGAAGAGAAAA
CAGAAGGGGAGGAAGAGCAGGAGGAGGAGGAGGAAGAGGAGGAAGAGGAAGGGGGAGACAAG
ATGACCAAGACAGGAAGCCACCCCAAACCTTGACCGAGAAGATCTT**TGA**CCCTTGCCTTTGAG
CCCCCAGGAGGGGAAGGGATCATGGAGAGCCCTCTAAAGCCTGCACTCTCCCTGCTCCACAG
CTTTCAGGGTGTGTTTATGAGTGACTCCACCCAAGCTTGTAGCTGTTCTCTCCCATCTAACC
TCAGGCAAGATCCTGGTGAAACAGCATGACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGG
TCCTGCTCCTAGAGATGAACTCTATÇCAGCCCCCTAATTGGCAGGTGTATGTGCTGACAGTA
CTGAAAGCTTTCCTCTTTAACTGATCCCACCCCCACCCAAAAGTCAGCAGTGGCACTGGAGC
TGTGGGCTTTGGGGAAGTCACTTAGCTCCTTAAGGTCTGTTTTTAGACCCTTCCAAGGAAGA
GGCCAGAACGGACATTCTCTGCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCA
GCAAACCGTGAAGGAGAATGGGACACTGGGTCATGGCCTGGAGTTGCTGATAATTTAGGTGG
GATAGATACTTGGTCTACTTAAGCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGT
GCTCAACTTTCTATATCGCTATTAACTTTTTTCTTTTTTTCTA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92256
><subunit 1 of 1, 248 aa, 1 stop
><MW: 28310, pI: 4.63, NX(S/T): 0
MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSKCEVCKLLSTELQAELSRTGRSREVL
ELGQVLDTGKRKRHVPYSVSETRLEEALENLCERILDYSVHAERKGSRLRYAKGQSQTMATLK
GLVQKGVKVDLGIPLLELWDEPSVEVITYLKKQCETMLEEFEDIVGDWYFHHQEQLQNLCEG
HVLPAEAETACLOETWTGKEITDGECKTEGEEEEEQEEEEEEEEEGGDKMTKTGSHPKLDREDL

Important features of the protein:

Signal peptide:

amino acids 1-21

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 106-110

N-myristoylation site.

amino acids 115-121

Amidation site.

amino acids 70-74

FIGURE 29

AAGTACTTGTGTCCGGGTGGTGGACTGGATTAGCTGCGGAGCCCTGGAAGCTGCCTGTCCTT
CTCCCTGTGCTTAACCAGAGGTGCCCATGGGTTGGACAATGAGGCTGGTCACAGCAGCACTG
TTACTGGGTCTCATGATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGA
GGCCCTCTTGGACGAGGACACCCTCTTTTGCCAGGGCCTTGAAGTTTTCTACCCAGAGTTGG
GGAACATTGGCTGCAAGGTGTTTCCTGATTGTAACAACTACAGACAGAAGATCACCTCCTGG
ATGGAGCCGATAGTCAAGTTCCCGGGGGCCGTGGACGGCGCAACCTATATCCTGGTGATGGT
GGATCCAGATGCCCCTAGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGGTAA
CAGATATCAAGGGCGCCGACCTGAAGAAAGGGAAGATTCAGGGCCAGGAGTTATCAGCCTAC
CAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAAGTTCTTTGTCTATCTTCA
GGAAGGAAAAGTCATCTCTCTCTTCCCAAGGAAAACAAAACCTCGAGGCTCTTGGAAAATGG
ACAGATTTCTGAACCGCTTCCACCTGGGCGAACCTGAAGCAAGCACCCAGTTCATGACCCAG
AACTACCAGGACTCACCAACCCTCCAGGCTCCCAGAGGAAGGGCCAGCGAGCCCAAGCACAA
AACCAGGCAGAGATAGCTGCCTGCTAGATAGCCGGCTTTGCCATCCGGGCATGTGGCCACAC
TGCTCACCACCGACGATGTGGGTATGGAACCCCTCTGGATACAGAACCCCTTCTTTTCCAA
ATTAAAAAAAAAATCATCAA

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92274
><subunit 1 of 1, 223 aa, 1 stop
><MW: 25402, pI: 8.14, NX(S/T): 1
MGWTMRLVTAALLLGLMMVVTGDEDENSPCAHEALLDEDTLFCQGLEVFYPELGNIGCKVVP
DCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHWLVTDIKGADLK
KGKIQQQELSAYQAPSPPAHSGFHRYQFFVYLQEGKVISLLPKENKTRGSKMDRFLNRFHL
GEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTRQR
```

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 169-173

Tyrosine kinase phosphorylation site.

amino acids 59-68

N-myristoylation sites.

amino acids 54-60, 83-89, 130-136

Phosphatidylethanolamine signature.

amino acids 113-157

FIGURE 31

GTCGACCCACGCGTCCGAAGCTGCTGGAGCCACGATTTCAGTCCCCTGGACTGTAGATAAAGA
CCCTTTCTTGCCAGGTGCTGAGACAACCACACT**ATG**AGAGGCACTCCAGGAGACGCTGATGG
TGGAGGAAGGGCCGTCTATCAATCAATCACTGTTGCTGTTATCACATGCAAGTATCCAGAGG
CTCTTGAGCAAGGCAGAGGGGATCCCATTTATTTGGGAATCCAGAATCCAGAAATGTGTTTG
TATTGTGAGAAGGTTGGAGAACAGCCCACATTGCAGCTAAAAGAGCAGAAGATCATGGATCT
GTATGGCCAACCCGAGCCCGTGAAACCCTTCCTTTTCTACCGTGCCAAGACTGGTAGGACCT
CCACCCTTGAGTCTGTGGCCTTCCCGGACTGGTTCATTGCCTCCTCCAAGAGAGACCAGCCC
ATCATTCTGACTTCAGAACTTGGGAAGTCATACAACACTGCCTTTGAATTAAATATAAATGA
CTGAACTCAGCCTAGAGGTGGCAGCTTGGTCTTTGTCTTAAAGTTTCTGGTTCCTCAATGTGT
TTTCGTCTACATTTTCTTAGTGTCATTTTCACGCTGGTGCTGAGACAGGAGCAAGGCTGCTG
TTATCATCTCATTTTATAATGAAGAAGAAGCAATTACTTCATAGCAACTGAAGAACAGGATG
TGGCCTCAGAAGCAGGAGAGCTGGGTGGTATAAGGCTGTCCTCTCAAGCTGGTGCTGTGTAG
GCCACAAGGCATCTGCATGAGTGACTTTAAGACTCAAAGACCAAACACTGAGCTTTCTTCTA
GGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGCATGACTAGCACA
GAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCATCCAGTCTTTAT
ATGTTGCCAATATACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAAGACCTTGTAACA
AAAATAATTCTTGGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGC
ATGACTAGCACAGAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCA
TCCAGTCTTTATATGTTGCCAATATACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAAG
ACCTTGTAACAATAATAATTCTTGTGTAAAGTTAAATCATTTTTGTCTAATTGTAATGTG
TAATCTTAAAGTTAAATAAACTTTGTGTATTTATATAATAATAAAGCTAAAACTGATATAAA
ATAAAGAAAGAGTAAACTG

FIGURE 32

MRGTPGDADGGGRAVYQSITVAVITCKYPEALEQGRGDPIYLGIONPEMCLYCEKVGEQPTL
QLKEQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIASSKRDQPIILTSELGKSY
NTAFELNIND

Signal sequence:

amino acids 1-17

N-myristoylation site.

amino acids 10-16

Cell attachment sequence.

amino acids 36-39

FIGURE 33

GCGAGGCTGCACCAGCGCCTGGCACC**ATG**AGGACGCCTGGGCCTCTGCCCCTGCTGCTGCTG
CTCCTGGCGGGAGCCCCCGCGCGCGGCCCACTCCCCGACCTGCTACTCCCGCATGCGGGC
CCTGAGCCAGGAGATCACCCGCGACTTCAACCTCCTGCAGGTCTCGGAGCCCTCGGAGCCAT
GTGTGAGATACCTGCCCAGGCTGTACCTGGACATACACAATTACTGTGTGCTGGACAAGCTG
CGGGACTTTGTGGCCTCGCCCCCGTGTTGGAAAGTGGCCCAGGTAGATTCCCTTGAAGGACAA
AGCACGGAAGCTGTACACCATCATGAACTCGTTCTGCAGGAGAGATTTGGTATTCCTGTTGG
ATGACTGCAATGCCTTGGAATACCCAATCCCAGTGACTACGGTCCTGCCAGATCGTCAGCGC
TAAGGGAACTGAGACCAGAGAAAGAACCCAAGAGAACTAAAGTTATGTCAGCTACCCAGACT
TAATGGGCCAGAGCCATGACCCTCACAGGTCTTGTGTTAGTTGTATCTGAAACTGTTATGTA
TCTCTCTACCTTCTGGAAAACAGGGCTGGTATTCCTACCCAGGAACCTCCTTTGAGCATAGA
GTTAGCAACCATGCTTCTCATTCCCTTGACTCATGTCTTGCCAGGATGGTTAGATACACAGC
ATGTTGATTTGGTCACTAAAAAGAAGAAAAGGACTAACAAGCTTCACTTTTATGAACAATA
TTTTGAGAACATGCACAATAGTATGTTTTTATTACTGGTTTAATGGAGTAATGGTACTTTTA
TTCTTTCTTGATAGAAACCTGCTTACATTTAACCAAGCTTCTATTATGCCTTTTTCTAACAC
AGACTTTCTTCACTGTCTTTCATTTAAAAAGAAATTAATGCTCTTAAGATATATATTTTACG
TAGTGCTGACAGGACCCACTCTTTCATTGAAAGGTGATGAAAATCAAATAAAGAATCTCTTC
ACATGGA

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA93011
><subunit 1 of 1, 136 aa, 1 stop
><MW: 15577, pI: 8.88, NX(S/T): 0
MRTPGPLPVLLLLLAGAPAARPTPPTCYSRMRALSQEITRDFNLLQVSEPSEPCVRYLPRLY
LDIHNYCVLDKLRDFVASPPCWKVAQVDSLKDKARKLYTIMNSFCRRDLVFLDDCNALEY
IPVTTVLPDRQR
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site.

amino acids 60-69

N-myristoylation site.

amino acids 16-22

FIGURE 35

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGA
CTTGACTCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCA
GCCGCCCAGTCCCGGCCCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCCTGTTTTTAC
TCCTCCTTTTTCATTCATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCA
AGCCGAGCGTGGAAGAATGGGGTTCTCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCC
CGATTCAAGCTTTCCCCAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTA
AGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAAC
ATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTTGATAACTTGAACCTGC
TAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCA
CTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAC TGATCGATGA
TTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATC
AACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTAT
GAAGAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTAT
CACAGAAAGCCAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCT
CAAAGGAAGCCAACAATTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAG
GCTGGAAAAATACCAGAGAAAGTGA CTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGG
AGAAAACGATGAAACAGTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTA
AAACCTACAGTGAAGACAACCTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTG
AAAAGTATTGATTCAGAAAAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAC
ACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTT
CCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAAT
GCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGA
CAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAA
AAGATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTAT
TTGGAAGCCATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGA
AGATTATGACCTTTCAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGA
AAGGCATCCTTGACAAGGAAGAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAA
TGGCAAAAGATCCAGGAGTCTTTCAACTGTTTCAGAAAACATAATATAGCTTAAAACACTTC
TAATTCTGTGATTAAATTTTTTGACCCAAGGGTTATTAGAAAGTGCTGAATTTACAGTAGT
TAACCTTTTACAAGTGGTTAAAACATAGCTTTCTTCCCGTAAAACTATCTGAAAGTAAAGT
TGATATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 36

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPEN
KPGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTK
SGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQA
HTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGK IPEKVTPMAAIQDGLAKGENDET
VSNTLTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFV
KMMVKYGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEE
AAKMEKEYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLS
KMRDFINKQADAYVEKGILDKEEAIAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

[illegible]

FIGURE 38

MALPPGPAALRHTLLLLPALLSSGWGELEPQIDGQTWAERALRENERHAFTCRVAGGPGTPR
LAWYLDGQLQEASTSRLLSVGGEAFSGGTSTFTVTAHRAQHELNCSLQDPRSGRSANASVIL
NVQFKPEIAQVGAKYQEAQGPGLLVVLFALVRANPPANVTWIDQDGPVTVNTSDFLVLDAQN
YPWLTNHTVQLQLRSLAHNLSVVATNDVGVTASLPAPGPSRHPSLISSDSNNLKLNNVRLP
RENMSLPSNLQLNDLTPDSRAVKPADRQMAQNNSRPELLDPEPGGLLTSQGFIRLPVLGYIY
RVSSVSSDEIWL

N-glycosylation sites:

amino acids 106-110, 119-123, 162-166, 175-179, 192-196, 205-209,
251-255, 280-284

Glycosaminoglycan attachment site:

amino acids 23-27

Casein kinase II phosphorylation sites:

amino acids 36-40, 108-112, 164-168, 282-286, 316-320

N-myristoylation sites:

amino acids 34-40, 89-95, 215-221, 292-298, 293-299

FIGURE 39

CGGGGACGGAAGCGGCCCCCTGGGCCCAGGGGGCTGGAGCCGGGCCGGGGCG**ATG**TGGAGCGC
GGGCCGCGGCGGGGCTGCCTGGCCGGTGCTGTTGGGGCTGCTGCTGGCGCTGTTAGTGCCGG
GCGGTGGTGCCGCCAAGACCGGTGCGGAGCTCGTGACCTGCGGGTCGGTGCTGAAGCTGCTC
AATACGCACCACCGCGTGCGGCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCA
GCAATCGGTGACCGGCGTAGAGGCGTCGGACGACGCCAATAGCTACTGGCGGATCCGCGGCG
GCTCGGAGGGCGGGTGCCCGCGCGGGTCCCCGGTGCGCTGCGGGCAGGCGGTGAGGCTCACG
CATGTGCTTACGGGCAAGAACCTGCACACGCACCACTTCCCGTCGCCGCTGTCCAACAACCA
GGAGGTGAGTGCCTTTGGGGAAGACGGCGAGGGCGACGACCTGGACCTATGGACAGTGCGCT
GCTCTGGACAGCACTGGGAGCGTGAGGCTGCTGTGCGCTTCCAGCATGTGGGCACCTCTGTG
TTCCTGTCAGTCACGGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGG
CATGCCCAGTGCCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTTCATCAAGCCTA
GTGTGGAGCCCTCTGCAGGTCACGATGAACTC**TGAG**TGTGTGGATGGATGGGTGGATGGAGG
GTGGCAGGTGGGGCGTCTGCAGGGCCACTCTTGGCAGAGACTTTGGGTTTGTAGGGGTCCTC
AAGTGCCTTTGTGATTAAAGAATGTTGGTCTATGAAA

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96857
><subunit 1 of 1, 221 aa, 1 stop
><MW: 23598, pI: 6.96, NX(S/T): 0
MWSAGRGGAAWPFVLLGLLLALLVPGGGAAKTGAELVTCGSVLKLLNTHHRVRLHSHDIKYGS
GSGQQSVTGVEASDDANSYWIRIRGGSEGGCPRGSPVRCGQAVRLTHVLTGKNLHTHHFPSPL
SNNQEVSAFGEDGEGLDLDLWTVRCGQHWEREAAVRFQHVGTSVFLSVTGEQYGSPIRGQH
EVHGMPSANTHNTWKAMEGIFIKPSVEPSAGHDEL
```

Important features of the protein:

Signal peptide:

amino acids 1-28

Glycosaminoglycan attachment site.

amino acids 62-66

N-myristoylation sites.

amino acids 16-22, 25-31, 27-33, 61-67, 71-77, 86-92, 87-93,
91-97, 190-196

Endoplasmic reticulum targeting sequence.

amino acids 218-223

Variable	Mean	SD	Min	Max	Median	Q1	Q3	Mode	Skewness	Kurtosis	Jarque-Bera	Prob > Chi-Sq
Age	35.5	10.5	20	65	35	30	40	35	0.1	3.0	0.99	0.99
Gender	0.5	0.5	0	1	0.5	0	1	0.5	0.0	0.0	0.99	0.99
Marital Status	0.7	0.5	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Education	12.5	2.0	8	16	12	11	13	12	0.1	3.0	0.99	0.99
Income	1500	500	500	3000	1500	1000	2000	1500	0.1	3.0	0.99	0.99
Health	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Stress	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Depression	0.4	0.5	0	1	0.4	0	1	0.4	0.0	0.0	0.99	0.99
Life Satisfaction	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Resilience	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Optimism	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Self-Efficacy	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Perceived Stress	0.5	0.5	0	1	0.5	0	1	0.5	0.0	0.0	0.99	0.99
Coping Strategies	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Emotional Stability	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Neuroticism	0.4	0.5	0	1	0.4	0	1	0.4	0.0	0.0	0.99	0.99
Extraversion	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Conscientiousness	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Agreeableness	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Openness	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Life Satisfaction (Scale 1-10)	7.5	2.0	5	10	7	6	8	7	0.1	3.0	0.99	0.99
Life Satisfaction (Scale 1-5)	3.5	1.0	2	5	3	2	4	3	0.1	3.0	0.99	0.99
Life Satisfaction (Scale 1-3)	2.0	0.8	1	3	2	1	3	2	0.1	3.0	0.99	0.99
Life Satisfaction (Scale 1-2)	1.5	0.5	1	2	1.5	1	2	1.5	0.0	0.0	0.99	0.99
Life Satisfaction (Scale 1-1)	1.0	0.2	1	1	1	1	1	1	0.0	0.0	0.99	0.99

Variable	Mean	SD	Min	Max	Median	Q1	Q3	Mode	Skewness	Kurtosis	Jarque-Bera	Prob > Chi-Sq
Age	35.5	10.5	20	65	35	30	40	35	0.1	3.0	0.99	0.99
Gender	0.5	0.5	0	1	0.5	0	1	0.5	0.0	0.0	0.99	0.99
Marital Status	0.7	0.5	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Education	12.5	2.0	8	16	12	11	13	12	0.1	3.0	0.99	0.99
Income	1500	500	500	3000	1500	1000	2000	1500	0.1	3.0	0.99	0.99
Health	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Stress	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Depression	0.4	0.5	0	1	0.4	0	1	0.4	0.0	0.0	0.99	0.99
Life Satisfaction	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Resilience	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Optimism	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Self-Efficacy	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Perceived Stress	0.5	0.5	0	1	0.5	0	1	0.5	0.0	0.0	0.99	0.99
Coping Strategies	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Emotional Stability	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Life Satisfaction	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Resilience	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Optimism	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99
Self-Efficacy	0.8	0.4	0	1	0.8	0	1	0.8	0.0	0.0	0.99	0.99
Perceived Stress	0.5	0.5	0	1	0.5	0	1	0.5	0.0	0.0	0.99	0.99
Coping Strategies	0.6	0.5	0	1	0.6	0	1	0.6	0.0	0.0	0.99	0.99
Emotional Stability	0.7	0.4	0	1	0.7	0	1	0.7	0.0	0.0	0.99	0.99

FIGURE 42

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96867

><subunit 1 of 1, 194 aa, 1 stop

><MW: 21431, pI: 8.57, NX(S/T): 0

MSALWLLLGLLALMDLSESSNWGCYGNIQSLDTPGASCGIGRRHGLNYCGVRASERLAEIDM
PYLLKYQPMMQTIGQKYCMDPAVIAGVLSRKSPGDKILVNMGDRTSMVQDPGSQAPTWISE
SQVSQTTEVLTTRIKEIQRRFPTWTPDQYLRGGLCAYSGGAGYVRSSQDLSCDFCNDVLARA
KYLKRHGF

Important features of the protein:

Signal peptide:

amino acids 1-19

N-myristoylation sites.

amino acids 23-29, 26-32, 35-41, 45-51, 50-56, 76-82, 156-162

Amidation site.

amino acids 40-44

FIGURE 43

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCC**ATG**AAAG
CCCTTATGCTGCTCACCCCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTCGCTGTCAC
TCCTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGAGCCAGG
ACAGCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAATCTGCGCT
GTGGCACACCAGAAGAGCCCTGTCAGGAGGCCTTCAACCAAACCAACCGCAAGCTGGGTCTG
ACATATAACACCACCTGCTGCAACAAGGACAACCTGCAACAGCGCAGGACCCCGGCCCACTCC
AGCCCTGGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTGGCTGCTGCAC**TGAG**
ACTCATTCCATTGGCTGCCCTCCTCCACCTGCCTTGGCCTGAGCCTCTCTCCCTGTGTCT
CTGTATCCCCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCATTTCTTTAATTAAACACTG
TTCCGAGTGGTCTCCTCATCCATCCTTCCACCTCACACCCTTCACTCTCCTTTTTCTGGGT
CCCTTCCCCTTCCCTTCCAGGACCTCCATTGGCTCCTAGAAGGGCTCCCCACTTTGCTTCCT
ATACTCTGCTGTCCCCTACTTGAGGAGGGATTGGGATCTGGGCCTGAAATGGGGCTTCTGTG
TTGTCCCCAGTGAAGGCTCCCACAAGGACCTGATGACCTCACTGTACAGAGCTGACTCCCCA
AACCAGGCTCCCATATGTACCCCATCCCCATACTCACCTCTTTCCATTTTGAGTAATAAA
TGTCTGAGTCTGGAAAAAAAAAAAAAAAAAAAA

FIGURE 44

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96878
><subunit 1 of 1, 125 aa, 1 stop
><MW: 13821, pI: 8.60, NX(S/T): 2
MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDRQSCRLEPGQQCLTTHAYLGKMWVFSN
LRCGTPEEPCQEAFNQTNRKLGITYNTTCCNKDNCNSAGPRPTPALGLVFLTSLAGLGLWLLH

Important features of the protein:

Signal peptide:

amino acids 1-18

N-glycosylation sites.

amino acids 77-81, 88-92

N-myristoylation site.

amino acids 84-90

Ly-6 / u-PAR domain protein signature.

amino acids 85-98

FIGURE 45

ACGGGCCGCGAGCGGCAGTGACGTAGGGTTGGCGCACGGATCCGTTGCGGCTGCAGCTCTGCA
GTCGGGCCGTTTCCTTCGCCGCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCGCGCT
ACCGCACCCAGGTTTCGGCCCCGTAGGCGTCTGGCAGCCCCGGCGCCATCTTCATCGAGCGCCAT
GGCCGCGAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCATTTGTTTC
TGCTGACCGCGGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCTGCGGGATGTA
AAAGCTCTTACCCTCCACTATGACCGCTATAACACCTCCCGCAGGCTGGATCCCATCCCACA
GTTGAAATGTGTTGGAGGCACAGCTGGTTGTGATTCTTATACCCCAAAGTCATACAGTGTC
AGAACAAAGGCTGGGATGGGTATGATGTACAGTGGAATGTAAGACGGACTTAGATATTGCA
TACAAATTTGGAAAACTGTGGTGAGCTGTGAAGGCTATGAGTCCTCTGAAGACCAGTATGT
ACTAAGAGGTTCTTGTGGCTTGGAGTATAATTTAGATTATACAGAACTTGGCCTGCAGAAAC
TGAAGGAGTCTGGAAGCAGCACGGCTTTGCCTCTTTCTCTGATTATTATTATAAGTGGTCC
TCGGCGGATTCCCTGTAACATGAGTGGATTGATTACCATCGTGGTACTCCTTGGGATCGCCTT
TGTAAGTCTATAAGCTGTTCCCTGAGTGACGGGCAGTATTCTCCTCCACCGTACTCTGAGTATC
CTCCATTTTCCCACCGTTACCAGAGATTCACCAACTCAGCAGGACCTCCTCCCCCAGGCTTT
AAGTCTGAGTTCACAGGACCACAGAATACTGGCCATGGTGCAACTTCTGGTTTTGGCAGTGC
TTTTACAGGACAACAAGGATATGAAAATTCAGGACCAGGGTTCTGGACAGGCTTGGGAACTG
GTGGAATACTAGGATATTTGTTTGGCAGCAATAGAGCGGCAACACCCTTCTCAGACTCGTGG
TACTACCCGTCCCTATCCTCCCTCCTACCCTGGCACGTGGAATAGGGCTTACTCACCCTTCA
TGGAGGCTCGGGCAGCTATTCGGTATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAG
GATATGGTGGTACCAGGAGACGATTAAAGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAAT
TTTGGATTTTTTCATCACTTTCTCTTTAGAAAAAAGTACTACCTGTTAACAATTGGGAAAAG
GGGATATTCAAAAGTTCTGTGGTGTATGTCCAGTGTAGCTTTTTGTATTCTATTATTTGAG
GCTAAAAGTTGATGTGTGACAAAATACTTATGTGTTGTATGTACAGTGTAACATGCAGATGTA
TATTGCAGTTTTTGAAGTGATCATTACTGTGGAATGCTAAAAATACATTAATTTCTAAAC
CTGTGATGCCCTAAGAAGCATTAGAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAA
AATTTAGTTTTAGGTGGTTGTAGCTGATGAGTTATTACCTCATAGAGACTATAATATTCTA
TTTGGTATTATATTATTTGATGTTTGCTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAA
TTATGCTAATTTGTGAGTTCTGATCACTTTTGAGCTCTGAAGCTTTGAATCATTAGTGGTG
GAGATGGCCTTCTGGTAACTGAATATTACCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTA
GAAGGTGTTGTGAATGACTCTGTGCTGGCAAAAATGCTTGAAACCTCTATATTTCTTTTCGT
TCATAAGAGGTAAAGGTCAAATTTTTCAACAAAAGTCTTTTAATAACAAAAGCATGCAGTTC
TCTGTGAAATCTCAAATATTGTTGTAATAGTCTGTTTCAATCTTAAAAAGAATCA

FIGURE 46

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96889

><subunit 1 of 1, 339 aa, 1 stop

><MW: 36975, pI: 7.85, NX(S/T): 1

MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLDPIP
QLKCVGGTAGCDSYTPKVIQCQNKGDGYDVQWECKTDLDIAYKFGKTVVSCEGYESSEDQY
VLRGSCGLEYNLDYTELGQLKLKESGKQHGFFASFSDDYKWKSSADSCNMSGGLITIVVLLGIA
FVVYKLFSLSDGQYSPPPYSEYPPFSHRYQRFNSAGPPPPGFKSEFTGPQNTGHGATSGFGS
AFTGQQGYENSGPGFWTGLGTGGILGYLFGSNRAATPFSDSWYYPSYPPSYPGTWNRAYSPL
HGGSGSYSVCSNSDTKTRTASGYGGTRRR

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 171-190

N-glycosylation site.

amino acids 172-176

Glycosaminoglycan attachment sites.

amino acids 244-248, 259-263, 331-335

Tyrosine kinase phosphorylation site.

amino acids 98-106

N-myristoylation sites.

amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,
270-276, 278-284, 312-318